Interactions between orientation and contrast modulations suggest limited cross-cue linkage

Schofield, Andrew; Yates, Tim

DOI:
10.1068/p5294

Document Version
Peer reviewed version

Citation for published version (Harvard):
Schofield, A & Yates, T 2005, 'Interactions between orientation and contrast modulations suggest limited cross-cue linkage', Perception, vol. 34, no. 7, pp. 769-792. https://doi.org/10.1068/p5294

Interactions between orientation and contrast modulations suggest limited cross-cue linkage

Andrew J. Schofield[1] and Timothy A. Yates[2]

School of Psychology
The University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK
Tel +44 (0)121 414 5644
Fax +44 (0)121 414 4234

1. Corresponding author: a.j.schofield@bham.ac.uk
2. t.yates@bham.ac.uk

5th November 2004

Abstract

Recent studies of texture segmentation and second-order vision have proposed very similar models for the detection of orientation and contrast modulations (OM and CM). From the similarity of the models it is tempting to assume that the two cues might be processed by a single generalized texture mechanism, however recent results (Kingdom, Prins Hayes, 2003) have suggested that these cues are detected independently or at least in a mechanism that is able to maintain an apparent independence between the cues. We test new combinations of OM and CM and find that CM at 0.4 c/deg facilitates the detection of OM at 0.2 c/deg when the peaks of contrast align with the extremes of orientation. There is also some evidence of weak facilitation of CM by OM under the same conditions. Further, this facilitation can be predicted by filter-rectify-filter channels optimized for the detection of each cue, adding weight to the argument that texture cues are processed in a single generalized mechanism that nonetheless achieves cue independence or near-independence in many circumstances. We also find that the amount of supra-threshold masking produced by an orientation cue depends on the overall percept formed by that cue.

1 Introduction

1.1 *Texture modulations*

In this paper we consider interactions between two types of texture cue (illustrated in Figs 1, 2 and 3); orientation modulations (OM) and contrast modulations (CM). These cues fall into the broad class of second-order stimuli as defined by Cavanagh and Mather (1989). Second-order stimuli - also called non-Fourier stimuli (Chubb and Sperling, 1988) - are those modulations that can not be detected by linear mechanisms, and whose impact on the Fourier amplitude spectrum of the image is obscure in the sense that no obvious spectral component is introduced at the spatial frequency and orientation of the modulation. The properties of the modulating signal can however be made accessible by the introduction of non-linear processing. Although OM and CM have each been classed as both texture cues and second-order modulations, they have seldom been studied in comparable paradigms.

Contrast modulations have not typically been used as stimuli in studies of texture segmentation but they have been used extensively to study motion processing (Chubb and Sperling, 1988; Wilson, Ferrera, and Yo, 1992; Derrington, Badcock, and Henning, 1993; Smith, Hess, and C.L.Baker, 1994; Solomon and Sperling, 1994; Solomon and Sperling, 1995; Ledgeway, 1994; Ledgeway and Smith, 1994; Ledgeway and Smith, 1995; Lu and Sperling, 1996; Smith and Ledgeway, 1997; Scott-Samuel and Georgeson, 1999). Where static contrast modulations have been studied this has typically been with the aim of elucidating the spatial properties of second-order vision in general, rather than in a texture segmentation type task (Sutter, Sperling, and Chubb, 1995; Schofield and Georgeson, 1999; Dakin and Mareschal, 2000; Schofield and Georgeson, 2000; Georgeson and Schofield, 2002; Schofield and Georgeson, 2003).

By contrast, OM has typically been used in the study of texture segmentation using non-sinusoidal, static modulations in a shape recognition type paradigm (Nothdurft, 1985; Landy and Bergen, 1991; Nothdurft, 1991; Nothdurft, 1993; Luschow and Nothdurft, 1993; Schofield and Foster, 1995) or to study the spatial properties of texture processing more directly, where sinusoidal modulations have sometimes been used (Keeble, Kingdom, Moulden, and Morgan, 1995; Kingdom, Keeble, and Moulden, 1995; Keeble, Kingdom, and Morgan, 1997; Gray and Regan, 1998) but not as a cue for motion processing.

Despite the fact that OM and CM have typically been studied within different experimental paradigms, suggested models for how these processes operate are strikingly similar. In both cases a two stage process has been suggested comprising two layers of linear spatio-temporal filters separated by a non-linear transducer, with multiple spatial channels at each stage (see for example Chubb and Sperling, 1988; Malik and Perona, 1990; Landy and Bergen, 1991; and Wilson et al, 1992). The first stage filters process the carrier signal (which may be oriented line elements, oriented Gabor patches, or anisotropic noise, for OM, and high frequency sinusoids, or various forms of spatial noise, for CM) while excluding other first-order signals from further processing by the second-order channel. The non-linearity then demodulates the second-order (non-Fourier) modulation signal so as to introduce Fourier components at the modulation frequency into the post-rectification 'image'. The second stage of filtering then detects the properties of the modulation while smoothing out any residual carrier components. The exclusion of non-carrier first-order components by the largely high-frequency first stage filters and the subsequent removal of the carrier by the largely low frequency second-stage filters ensures the independence of the first and second-order mechanisms. Such independence has been measured empirically (Schofield and Georgeson, 1999; Georgeson and Schofield, 2002). In common with the majority of the second-order literature we call this class of models filter-rectify-filter (FRF) models.

FRF models for texture segmentation based on OM and the detection of moving CM stimuli differ in terms of the final processing of the second-stage outputs (e.g. object shape encoding vs. motion encoding) and in the proposed connectivity between the first and second-stage channels. Texture models based on orientation tend to compute the difference between first-stage filters (falling within the 'receptive fields' of the second-stage filters) so as to compute opponent orientation signals (see for example, Landy and Bergen, 1991, but see also Motoyoshi and Kingdom, 2003), whereas for a CM signal it is more efficient to either sum across orientation channels or to keep the sub-channels separate (see for example Schofield, 2000a).

The similarity of the FRF models for OM and CM detection suggests that these two cues might be processed by a single generalized texture mechanism. This would be an efficient processing strategy compared to multiple FRF-like mechanisms for each cue type. Kingdom, Prins, and Hayes (2003) tested for cue independence using a cross-cue facilitation and masking paradigm (*cf* Legge and Foley, 1980). In this method thresholds for detecting a test cue presented in one of two temporal intervals is measured as a function of the strength of a pedestal cue that is presented in both intervals. If the two cues interact via an expansive non-linear mechanism then the presence of a weak (peri-threshold) pedestal signal will increase sensitivity to the test cue when the two are in phase and reduce sensitivity when they are in anti-phase. If the cues interact via a later masking process then high strength backgrounds will reduce sensitivity to the test signal. Kingdom et al (2003) found no facilitatory interactions between any combination of orientation, contrast, or spatial frequency (FM) modulations. They did however verify that each cue is capable of inducing facilitation on itself.

Although they did not draw this conclusion Kingdom et al's (2003) lack of cross cue interaction suggests that texture cues are detected in entirely independent mechanisms. However, as Kingdom et al (2003) conclude the result is also consistent with a single multi-channeled mechanism that achieves apparent independence due to a combination of off-axis processing and contrast normalization. This finding rests on the idea that the best channel for detecting variations about a carrier signal is sometimes not the channel that is tuned to the dc (or mean) properties of the carrier. Therefore modulations about the same carrier might be detected within a single mechanism comprising multiple sub-channels with each cue being detected by a different sub-channel even when they are presented on the same carrier. Direct evidence in support of this off-axis idea was provided by Prins and Kingdom (2002) who showed that the responses of second-order mechanisms can be altered after adapting to a pure luminance sinusoid, and that the best such adapter seldom shares all of its properties with the carrier but would stimulate the predicted off-frequency or off-orientation channels well. Off-axis processing alone still predicts facilitation of OM and FM by CM but this was not found experimentally, leading Kingdom et al (2003) to propose a further contrast normalization process (*cf* Landy and Bergen, 1991) that reduces the impact of CM pedestals on OM and FM test cues.

In essence Kingdom et al (2003) found an interesting but null result and replicated this result with a model that showed no trends. The problem with this approach is that the data could equally well be modeled by a bank of entirely separate FRF mechanisms. (The adaptation results of Prins and Kingdom, 2002, do not exclude the multiple mechanisms model). Further, the lack of interaction found by Kingdom et al (2003) might be true only of the conditions that they tested. This is an important issue as the existence of multiple separate multi-channeled texture mechanisms would make human vision considerably more complex than would be the case if processing takes place in a single multi-channeled mechanism.

We now extend Kingdom et al's (2003) results to new conditions. We concentrate on stimuli where the CM cue has twice the spatial frequency of the OM cue. In the relatively few cases where our test conditions overlap our findings agree largely with those of Kingdom et al (2003), however we do find limited cross cue facilitation in the case where the contrast peaks of CM align with the extremes (peaks and troughs) of orientation in the OM cue. Kingdom et al (2003) did not test this configuration. Our interactions can be predicted by a FRF model similar to that proposed by Kingdom et al (2003) but without contrast normalization. We thus produce positive evidence in favor of a generalized texture mechanism with nearly-independent sub-channels but some cross cue linkage. Before outlining our experimental aims in more detail we now discus the various perceptual configurations that can be formed by OM and CM stimuli.

1.2 *Perceptual properties of texture modulated stimuli*

1.2.1 *Orientation modulations.*

When the orientation of anisotropic noise is modulated by a sinusoidal modulation pattern a number of different percepts can be obtained. If the orientation of the modulation is orthogonal to the dc-orientation (alternatively, mean orientation or zero-crossing orientation) of the carrier then the texture pattern forms a corrugated, or depthy percept (Fig. 1a). When modulation orientation and dc-orientation are parallel a 'herring-bone' percept is observed. This is generally seen as being flat and is also suggestive of segmented regions of texture (Fig. 1c). A key difference between these stimuli is that in orthogonal pairings the contours formed by the texture elements tend towards being continuous across the image where as in the parallel configurations they are not continuous. We therefore call the orthogonal parings 'continuous' and the parallel pairings 'segmented'.

Orientation modulated stimuli have other interesting perceptual properties. The modulation frequency used in Figures 1a, and c is about 1.5 cycles per image. Figures 1b and d show cross sections of the modulator signals for figures 1a and c respectively. Consider Fig. 1a and b, if we define positive orientations as being clockwise from vertical the peaks of the modulation occur when the sinewave modulator is +1 corresponding to a peak orientation of about 115°, the troughs of modulation occur when the sinewave has the value -1 corresponding to an orientation of 65° (giving a peak-to-peak modulation amplitude of 50°). The dc-orientation (90°) is found at the zero-crossings of the modulator and occurs twice in each cycle of modulation. For the continuous percept the signal has the appearance of a sinusoid that is phase shifted relative to the underlying modulation. (We use the ° symbol to indicate degrees of rotation and deg to indicate degrees of visual angle). In the segmented case (Fig. 1c and d) the dc-orientation (here 0°, vertical) again occurs twice per cycle of modulation. We note that a filter like mechanism tuned to the dc-orientation of an OM stimulus would produce a frequency doubled (though non-sinusoidal) output in response to the images of Fig 1a and c.

Figure 1 about here

1.2.2 *Contrast modulations.* Figure 2a shows a contrast modulated stimulus with its associated modulator signal (Fig 2b). A filter mechanism tuned to the dc-orientation for this stimulus would produce its maximum output only at the peaks of modulation - that is once per cycle of the modulation. So in order to equate the spatial frequencies in the output of a filter mechanism tuned to the dc-orientation we have to produce CM modulations that are twice the spatial frequency of our OM signal, as in Figs 2 and 3. We note that the outputs of filters tuned to other orientations will be dominated by a component at the OM modulation frequency, however, given that Kingdom et al (2003) concentrated on interactions between

OM and CM at the same spatial frequency, we now investigate the frequency doubled case. The perceptual qualities of CM stimuli are rather vague. CM does not produce either a strong depth or segmentation percept although we note that in general they seem quite flat and that regions which differ considerably in contrast can be seen as a change in material or the overlaying of a semi-transparent 'surface' onto the texture (not unlike patchy fog).

Figure 2 about here

1.3 *Experimental aims*

Although Kingdom et al (2003) undertook a fairly extensive study of the interactions between OM, CM and spatial frequency modulations they only considered orientation modulations that formed a continuous percept (see section 1.2 and Fig 1a). As we outline above, OM can be configured so as to produce a range of percepts and we postulate that such differences might affect facilitation and masking. We thus test for interactions when OM formed both the segmented and continuous percepts described above. Further, although Kingdom et al (2003) included some stimuli in which the spatial frequency of the CM cue was doubled relative to the OM, they did not test the frequency-doubled CM case in as much detail as the non frequency-doubled case. In particular they only varied the relative phase of OM and CM in the non frequency-doubled case and then only tested with pedestal stimuli presented at their own detection threshold. Here we test frequency-doubled CM in both 'in-phase' and 'out-of-phase' combinations with a range of pedestal strengths and hypothesize that facilitation will be found in one phase relationship but not the other. If this hypothesis holds it will provide strong evidence that OM and CM are detected within a single multi-channeled mechanism.

2 General methods

2.1 *Experimental procedure*

Our experimental procedure was based upon the facilitation and masking or 'dipper' paradigm (Legge and Foley, 1980) with one cue used as a pedestal or background against which the other cue had to be detected. In all trials, participants were asked to indicate which of two temporal intervals contained the test cue (two interval forced choice, 2ifc). We routinely tested baseline conditions in which there was no pedestal cue alongside the pedestal trials. These baseline thresholds were used as the basis for assessing the level of facilitation or masking present in a given condition.

Before a dipper experiment proper can begin it is necessary to obtain an initial estimate of the observers threshold for the pedestal cue presented alone. Here we either based our estimates on the no pedestal conditions of a previous experiment in which the new pedestal cue had acted as test or we estimated the threshold for the new pedestal in a separate 2ifc experiment. These initial estimates were used only to determine an appropriate range of modulation strengths for the pedestal cue, not as baselines for assessing the degree of facilitation.

Detection thresholds for the test cue were typically recorded against several levels of pedestal cue as described in the text for each experiment. Several pedestal levels were grouped for presentation within each session. Each condition was tested in at least two experimental sessions. Allocation of pedestal levels to sessions was done on a random basis (no two sessions contained the same set of test levels). A no-pedestal condition was presented within each session. Within a session pedestal levels were presented in blocks of 10 trials. Block selection was essentially random but when staircases were used blocks were selected such that all of the staircases completed at about the same time. The grouping

of test levels into sessions, the random allocation of levels to sessions, and the use of sub-blocks within sessions was designed to minimize order effects while controlling observer uncertainty in relation to the task. Both staircase and constant stimulus methods were used as described in the specific methods sections for each experiment.

2.2 *Stimulus timing and positional jitter*

The total duration of a presentation interval was 706ms and the contrast of the whole stimulus was smoothed on and off by half a cycle of a raised cosine lasting 141ms. For most experiments the absolute phase of the modulation, and the noise sample used changed once every 53ms (every 8th monitor refresh). This temporal jitter introduced a dynamic element into the stimuli, but no coherent motion. The jitter prevented observers from using features in the pedestal as positional cues to reduce their uncertainty of the phase of the test cue which might otherwise have resulted in supra-threshold, phase independent, facilitation. The inter-stimulus interval was 706ms. A fixation cross was displayed before and after but not during each stimulus interval. In order to test our jittered presentation method we tested for facilitation of a luminance test by a luminance pedestal. We found the classic dipper profile suggesting that jittering the position of the stimuli does not of itself disrupt facilitatory interactions (data not shown).

2.3 *Stimulus construction*

2.3.1 *Orientation modulations.* Extracted regions of example stimuli containing only orientation modulations are depicted in Fig 1. OM stimuli were formed by placing Gabor patches (64x64 pixels) onto the image with one patch centred on each pixel (when Gabors overlaped the edge of the image they were cropped but see section 2.3.4). Gabors were added arithmetically and when they overlapped their amplitudes (not dc-luminances) were added. In common with Kingdom et al (2003) the pixel values in the image were clipped at ±3 standard deviations from the mean so as to boost rms contrast. This clipping introduced

some additional frequency components into the stimuli but had minimal effect on their appearance. Gabors were generated in the space domain but with space constants chosen so as to achieve the desired frequency (1.22 octaves, full-width at half-height) and orientation (14.4°, half-width at half-height) bandwidths. These bandwidths matched estimates of physiological and psychophysical channel bandwidths at the Gabors' central frequency of 8c/deg (DeValois, Albrecht and Thorell, 1982; Snowden, 1992). One patch was made for each orientation in the range 0-180° at 0.1° intervals. The amplitude of the Gabor patch centred on location *x,y* in the stimulus was determined by the value of the corresponding pixel in a white noise image. The orientation of the patch at *x,y* in the stimulus was determined by adding the dc-orientation of the carrier, θ_c, to the value of the corresponding pixel in a modulator image containing a low frequency sinusoid (frequency f_{om}, orientation θ_{om}, and phase φ_{om}) with amplitude A_{om}. That is,

$$\theta_g(x,y) = \theta_c + A_{om}\left\{\cos\left[2\pi f_{om}\left[x\cos(\theta_{om}) + y\sin(\theta_{om})\right] + \varphi_{om}\right]\right\} \qquad \text{Eq 1}$$

2.3.2 *Contrast modulations.* Figure 2 shows an extract from a stimulus containing CM only; stimuli containing both CM and OM are depicted in Fig 3. Contrast modulations were imposed onto the orientation-modulated Gabor noise images according to equation 2:

$$I(x,y) = nO(x,y)\cdot\left(1 + m\cos\left\{2\pi f_{cm}\left[x\cos(\theta_{cm}) + y\sin(\theta_{cm})\right] + \varphi_{cm}\right\}\right) \qquad \text{Eq 2}$$

where *m* is the modulation depth of the CM signal (a sinusoid of frequency f_{cm} orientation $\theta_{cm,}$ and phase φ_{cm}), $O(x,y)$ is an orientation modulated stimulus (as described in section 2.3.1) with contrast *n*. Equation 2 can be re-written as:

$$I(x,y) = nO(x,y) + mnM(x,y)O(x,y) \qquad \text{Eq 3}$$

where $M(x,y) = \cos\{2\pi f_{cm}[x\cos(\theta_{cm}) + y\sin(\theta_{cm})]\varphi_{cm}\}$. This equation has the advantage of separating the OM only term $nO(x,y)$ from the side-band or beat term $mnM(x,y)O(x,y)$ which conveys the information content associated with the contrast modulations. In some of the experiments described stimuli were constructed according to equation 2 with *n* set such that the final rms contrast of an OM only image was, on average, 0.15 with 'mid-grey' pixels in the image presented at the mean luminance of the monitor. We call this the single frame method. In other experiments images were generated according to equation 3 with the two images being presented (at contrasts 2*n* and 2*mn*) in alternate frames of the video sequence such that the time averaged contrasts of each component were *n* and *nm* respectively. In this frame interleave method the addition of the component images to form the final stimulus took place in the eye, the frame rate being well above the flicker fusion frequency. Paired images in the frame interleave signal were based on the same orientation signal ($O(x,y)$). The carrier contrast, given by *n* was held constant at 0.15 while the modulation depth of the CM signal (*m*) was allowed to vary (see Schofield and Georgeson, 1999, for a more complete description of the frame interleave method).

With appropriate settings of the OM amplitude and CM modulation depth it was possible to generate stimuli with no modulation at all (i.e. uniform-contrast, uniform-orientation anisotropic Gabor noise), orientation modulations alone (Figs 1a,c), contrast modulations alone (Fig 2a) or both contrast and orientation modulations (Fig 3a,b). Note that in common with Kingdom et al (2003) when OM and CM were presented together they were applied to the same carrier signal rather than being generated as completely separate images that were later added. Images were generated for four types of OM modulation providing both segmented and continuous percepts for vertical and horizontal dc-orientations. OM and CM were combined in two phase relationships: peaks of contrast align with orientation zero-

crossings (phase relationship A, Fig 3d) and peaks of contrast align with extreme orientations (phase relationship B, Fig 3a).

Figure 3 about here

2.3.3 *Modulation phase and noise samples.* For every condition and pedestal level tested we generated a bank of stimuli with 10 different absolute phases, generated from two different white noise samples (20 stimuli in all). Fresh stimuli were used for each observer. No two stimuli in our experiments were generated from the same white noise sample except for carrier side-band pairs when the frame interleave method was used.

2.3.4 *Raised cosine window.* Prior to presentation all stimuli were windowed by a circular raised cosine spatial window. Stimuli had full contrast over a central region of radius 6 deg, outside this region stimulus contrast tapered smoothly according one half period of a raised cosine such that contrast fell from maximum to zero across a 1.0 deg wide annulus. Outside of the window stimuli had mean luminance to the limit of the monitors display area.

2.4 *Equipment*

Images were generated off-line on a Pentium 4 PC (HiGrade Ltd, UK) using the MatLab computer program (Mathworks Ltd, UK) with the image processing toolbox and additional functions, and were presented on a high resolution 21" colour monitor (Sony Multiscan CPD-GM520, Sony Inc, Japan – mean luminance 53cdm^{-2}) using a VSG2/5 graphics card with extended memory (Cambridge Research Systems Ltd, UK) under the control of the PC. Images were displayed in monochrome (grey) at a frame rate of 170Hz (85Hz per composite image when frame interleaving was used). Images were displayed centrally within a square region of the monitor screen corresponding to 512x512 pixels (side length = 25cm or 14.25 deg at the viewing distance of 1m).

2.5 *Calibration*

Careful calibration is required when testing second-order vision (both CM and OM) as non-linearities in the display equipment can induce luminance artefacts in second-order stimuli. The monitor's gamma non-linearity was corrected using look-up tables in the VSG to provide gamma corrected stimuli with pseudo 15 bit luminance resolution. The appropriate correction was determined from the relationship between pixel value and screen luminance obtained using a ColourCal (CRS Ltd, UK) digital luminance meter interfaced to the computer and using CRS's proprietary calibration routines. This calibration was confirmed using a Minolta LS110 luminance meter (Minolta Inc, Japan) to check linearity after the CRS calibration had been applied. Linearity was rechecked periodically.

A more serious problem is presented by the adjacent pixel non-linearity (APNL; Klein, Hu and Carney, 1996; Mulligan and Stone, 1989) which results in a systematic reduction in luminance with increased image contrast, and is worst when stimuli contain high frequency components that vary across the display raster. APNL thus introduces contrast and orientation dependent luminance artifacts associated with high frequency image components. APNL cannot be corrected by a standard gamma-correction procedure but we minimised it by choosing a mid-range Gabor carrier frequency that is unlikely to present a problem. We also used a high bandwidth monitor (200MHz) and set its brightness and luminance controls to minimise gamma prior to gamma correction (uncorrected gamma ≈ 2.0). As a check, we compared the luminance of key parts of example stimuli and found variations consistent only with random sampling of the noise and normal variations with spatial position across the monitor. Finally we blurred example images from our experiments using a tracing paper screen placed in front of the monitor. This rendered the Gabor noise and hence both OM and CM invisible but would not have removed low frequency luminance cues. We did not observe any luminance variations under these conditions.

2.6 *Observers*

The two authors and one experienced psychophysical observer, who was not aware of the purpose of the experiments, took part in experiment 1. Only the two authors took part in experiment 2. All had normal or corrected-to-normal visual acuity and either negligible or corrected astigmatism.

2.7 *Data processing and threshold estimation*

Regardless of the data collection method all the data for a given pedestal level, stimulus configuration, and observer were fit with a Weibull function from which 82% thresholds were extracted using a bootstrap procedure (Foster and Bischof, 1991). Data from the two dc-orientations contributing to each percept were combined prior to threshold estimation. Curve fits and bootstrap procedures were executed in Matlab using the Psychofit (now Psignifit) routines available at http://bootstrap-software.org/psignifit/ as described by Wichmann and Hill (2001a,b).

3 Experiment 1: OM Pedestal CM test

3.1 *Procedural detail*

In this experiment the observers had to detect contrast modulations against an OM pedestal. Individual thresholds for detecting OM alone were first measured using a 3 up 1 down dual-staircase procedure that varied the (zero to peak) amplitude of the OM in 0.5° steps. The recorded thresholds are shown in Table 1 and are consistent with those reported previously (Kingdom et al, 1995). Thresholds for detecting CM at twice the frequency of the OM modulator were then measured in the presence of OM using staircases that varied modulation depth in 1 dB steps relative to a modulation depth of 0.01 (signal strength in dB = $20.\log_{10}(s/0.01)$ where s is the linear signal strength). Images were presented using the

frame interleave method. The OM pedestal could take one of 7 levels at the following multiples of its own detection threshold: 0.25, 0.5, 1, 2, 4, 16, plus 0.0 for the no pedestal condition. All four configurations of OM dc-orientation and modulation orientation, and two relative phases (phase relationship A, peaks of contrast align with orientation zero-crossings; phase relationship B, peaks of contrast align with peaks and troughs of orientation), were tested making a total of 8 stimulus configurations but results were later collapsed across dc-orientation. Trials were grouped into sessions such that only one dc - modulator orientation pair (e.g. vertical dc-orientation, horizontal modulation) was tested in any one session. Three levels of background signal and both relative phases were tested in each session with pedestal levels assigned to sessions on a random basis. Two staircases were run per pedestal level per session. CM detection with no OM pedestal was measured by two staircases in *every* session (relative phase is not relevant in this no pedestal case).

Table 1 about here

3.2 *Results*

The results for experiment 1 are shown in figure 4. The x-axes represent the OM pedestal level normalized to each observers' OM only threshold. The y-axes are normalized relative to each observers' CM threshold in the absence of OM. Deviations below the line y=1 would indicate facilitation, deviations above it masking. Error bars on plots for individual observers represent 95% confidence intervals that are 1.96 times the standard errors estimated by the bootstrap threshold analysis but adjusted to take account of the normalization process using the formula,

$$\sigma_{i(norm)} = \sqrt{\left(\sigma_0 y_i / y_0^2\right)^2 + \left(\sigma_i / y_0\right)^2} \ , \qquad \text{Eq 4}$$

where σ_0 is the standard error for the baseline condition, σ_i the un-normalized standard error for the ith pedestal level, and y_0 is the baseline threshold. This formula is derived from the general rule that when $y = f(a,b)$, $\sigma = \sqrt{\sigma_a \, \partial f / \partial a + \sigma_b \, \partial f / \partial b}$ with f in our case

given by $y_{i(norm)} = y_i / y_0$, y_i being the un-normalized threshold at the *i*th pedestal level. This formula ensures that the error inherent in the estimation of the baseline thresholds is appropriately transferred to the post normalization results. Figure 4d and h show the results averaged across all observers and sub-conditions (note this is the mean of the individually normalized results, error bars represent 95% confidence intervals on this mean calculated in the normal way). The dashed and dotted lines in each panel represent predictions from the computational models described in section 6.2. Note these models had no free parameters, therefore models were not fit to individual data and hence vertically aligned panels show the *same* model predictions.

Figure 4 about here

The data from Experiment 1 are quite noisy and variations from the no pedestal threshold are small. There are, however, some trends (strongest for AJS but present for all observers) suggested in the data that might represent the facilitatory interactions predicted by our hypothesis. When the OM variations form a segmented percept and OM and CM are presented in phase relationship B (peaks of contrast align with peaks and troughs of orientation) there is a very weak, slightly late dip followed by some masking. In the continuous-percept, phase relationship B case we see a general if slight decrease in threshold with pedestal strength, this facilitation saturates but there is no real masking. The significance of these trends for individual observers can be assessed from the 95% confidence intervals; all observers have a threshold minimum (at around 2-4 times OM threshold) whose range does not contain 1.0, with 95% confidence (but see also section 6.1.2). A non-linear (quadratic) regression conducted on the combined data was significant (see Section 6.1.1) suggesting that the facilitation is genuine. The difference between the segmented and continuous cases in the 16 times threshold pedestal condition was significant (t=7.055, df=2, p<0.05) suggesting that segmented OM masks CM whereas

continuous OM does not. In the phase relationship A condition (peaks of contrast align with dc-orientation) we see no reliable facilitation although there is some late masking which has statistical support (see section 6.1). We can be confident that the facilitatory effects observed depend on the phase relationship between the test and pedestal.

4 Experiment 2: CM background OM test

4.1 *Procedural detail*

In this experiment the observers had to detect OM against CM pedestals. Initial threshold estimates for the CM pedestals were taken from Experiment 1, (see Table 1) and are consistent with previously reported threshold modulation depths recorded in a variety of noise types (Schofield and Georgeson, 2003). The CM pedestal could take one of 6 levels at the following multiples of its own detection threshold: 0.25, 0.5, 1, 2, 4, plus 0.0 for the no pedestal case. Data were collected using the method of constant stimuli and the single image combination method. Only the two authors took part in this experiment. All other procedural details were identical to those of Experiment 1.

4.2 *Results*

The results for experiment 2 are shown in figure 5. The data are presented in the format used for Figure 4 but note the altered axes and that the model curves represent different models.

Figure 5 about here

Again the data have only weak trends but these are consistent enough to suggest that there is a genuine, phase dependent, facilitatory interaction. Whenever the stimuli were presented in phase relationship B (peaks of contrast align with peaks and troughs of orientation) we

found a general reduction in threshold as pedestal strength increased. This facilitation began before the CM pedestal reached its own threshold but saturated at twice this threshold and was statistically significant for the combined data (see Section 6.1.1). The significance of the trends for individual observers can be assessed from there 95% confidence intervals (but see also section 6.1.2). Combined results for phase relationship A showed masking only but no facilitation and this result has some statistical support (see section 6.1). We note that there is considerable inter-subject and inter-condition variability in phase relationship A (which we think may result from poor estimates of the no pedestal thresholds against which all data are normalized) but that each observer-condition pairing shows the same pattern – initially flat with a late rise. The facilitation found in this experiment is dependent on relative phase, thus ruling out the action of a simple positional cue. We found no reliable differences between the segmented and continuous percepts.

5 Control experiment, within cue interactions

Given the lack of strong cross cue facilitation it is pertinent to ask if CM and OM cues produce strong dipper like interactions on themselves. Dippers for CM on CM have been reported previously (Schofield and Georgeson, 1999; Kingdom et al, 2003) and reveal clear, relatively deep, peri-threshold facilitation. Thus we conclude that CM can facilitate itself verifying that this mechanism susceptible to facilitation by suitable threshold-level pedestals. Kingdom et al, (2003) report peri-threshold facilitation of OM on itself but we thought it wise to confirm this result. We thus tested OM thresholds in the presence of OM pedestals. Stimuli were not phase jittered in this control experiment but the presentation time was reduced to 200ms. One author (TAY) took part in the experiment. The results are shown in figure 6. Axes are scaled relative to the no-pedestal OM threshold. We found a strong peri-threshold dip that starts when the pedestal is sub-threshold. Thresholds return to baseline but there is no late masking, this result is consistent with previous data.

Figure 6 about here

6 Statistical analyses and modeling

When experimental data show only weak trends it is useful to verify their statistical significance. As some of the analyses we have used are complex we devote a separate section to them here (section 6.1). It is also helpful to model such data sets with a computational model that is (at least to the first approximation) biologically plausible we present such models in section 6.2.

6.1 *Statistical analysis*.

6.1.1 *Averaged data*. We combined the data from all observers, collapsing across OM percept, in order to obtain threshold estimates for the average observer (see solid lines in figs 4d,h and 5c,f). Treating the data from individual observers as independent observation of these averaged values we then have sufficient data to contemplate the calculation of significance for a regression fit (this would not be sensible for individual data – see section 6.1.2). We fit non-linear regression curves with quadratic models to our combined, normalized data ignoring the points at zero pedestal as they were fixed at 1.0 and thus not free to vary. The following quadratic fits were significant: OM-test vs CM, phase relationship B (Rsq=0.765, F=12.0, p=0.0006, df=2,17), CM-test vs OM, phase relationship B (Rsq=0.599, F=9.22, p=0.0007, df=2,33), and CM-test vs OM, phase relationship A (Rsq=0.57, F=7.922, p=0.0015, df=2,33). Although the quadratic fit for OM-test vs CM, phase relationship A was not significant a linear fit to these data was (Rsq=0.479, F=5.364, p=0.0325, df=1,18). These results remain significant when corrected for multiple comparisons using the Larzelere and Mulaik method assuming 4 regression comparisons (see Howell 1997). Detailed analyses of the significance of each coefficient in the regression fits revealed that while significant overall the quadratic term for CM-test vs OM,

phase relationship A is not required, a linear fit would be about as good as the quadratic one in this case. Thus for both test cues phase relationship B is best modeled by a quadratic curve reproducing both facilitation and masking whereas the phase relationship A cases are adequately modeled with linear (masking only) functions.

6.1.2 *Individual data*. Because we estimated threshold levels from Weibul fits to all the available data for a given condition we have only one threshold estimate per observer per condition per level. The 95% confidence intervals shown on the individual observers' graphs were estimated by bootstrap procedures. Thus, there are no traditional statistics (such as regression) that can be calculated, meaningfully, on our data for individual observers other than to note the cases of individual data points whose confidence intervals do not contain 1.0. Such, points can be regarded as evidence in favor of significant facilitation or masking in the individual data sets.

We can, however, take the analysis of individual data a little further by using Monte Carlo simulations. We found the best-fit quadratic polynomial for each curve figs 4 and 5 (except the averaged data and models) and compared predicted thresholds for each fit to simulated data sets generated randomly with each data point selected from a normal distribution with the correct (constrained by the data) mean and standard deviation. We then calculated the sum of squared errors for each simulated data set. We repeated this process 10,000 times for each comparison, *without* redoing the quadratic fit. We then repeated this process for comparisons between simulated data sets and a 'curve' representing the null hypothesis that the pedestal cue has no effect (that this the line y=1.0). We then counted the number of times any of the 10,000 sum of squared errors for the quadratic fit exceeded any of the 10,000 estimates for the null curve. This count divided by the total number of comparisons yields an estimate of the probability that the null hypothesis is a better fit to the data than

the single best fit quadratic to the initial threshold estimates. Table 2 shows the probability estimates thus generated.

Table 2 about here

It should be noted that if two curve fits provide equivalently good descriptions of the data then the probability estimate that one fit is better than the other will yield a value close to 0.5. Probability estimates of 0.05 or 0.95 indicate that one model is significantly better than the other depending on the order of the comparison made. In this case $p<0.05$ would indicate that the quadratic fit is significantly better than the null and $p>0.95$ would indicate that the null hypothesis is the better fit.

In general the quadratic fits are significant better than the null curve for all observers / conditions where one would expect a significant difference based on the error bars of figs 4 and 5. We note particularly that the quadratic fits are significantly better for phase relationship B, for both test cues, except for TAY in the segmented percept, OM test case. For phase relationship A the quadratic fits are always better than the null curve (in many cases significantly so or approaching significance) but in general the quadratic fits fair less well in this condition suggesting that the data are rather flat for phase relationship A. These individual significance estimates are best seen as moderating conclusions based on the significance of the averaged data. We conclude that, on the whole, the Phase relationship B data deviate significantly from the null hypothesis for the averaged data and (mostly) for individual data also, whereas for phase relationship A the deviation is significant only for the averaged data and might therefore be regarded as only marginally significant overall.

6.2 *Computational models*

6.2.1 *General notes.* We constructed a series of simple computational models in an attempt to predict the data in each cue combination and also to test some specific ideas as

outlined below. The models had no truly free parameters and hence they were *not* optimised to fit any data set as such. The models were based on the FRF principle and were similar to those used by Landy and Bergen (1991), Schofield (2000a), and Kingdom et al (2003). Individual models were intended to capture only the operation of one part of a generalized texture processing mechanism; we do not present these models as complete texture processors. The models are doubtless a gross over simplification of the texture mechanisms in the human visual system (HVS). However, while we remain neutral as to the biological implementation of texture processing in the HVS we find models such as ours informative, and, with a good choice of filter (channel) bandwidths, we believe them to have some biological validity as models of steady state processing.

Images were subjected to two stages of filtering (all filtering was conducted in the Fourier domain). Each first stage filter was implemented as a quadrature-phase pair of Gabor filters whose outputs were squared and summed to produce an 'energy' map selective to the properties of the filters. We thus introduced a demodulating non-linearity via the square law transducer function at the output of the first stage filters. All first stage filters had a preferred spatial frequency equal to that of our carrier Gabors and a frequency bandwidth of 1.22 octaves and an orientation bandwidth of 14.4° these being both optimal for the stimuli and biologically plausible. These bandwidths represent *average* estimates for human channels (Snowden, 1992) and cells in area V1 of monkey (DeValois, Albrecht and Thorell, 1982).

The second stage filters consisted of a single quadrature-phase Gabor pair tuned to the spatial frequency and orientation of the texture modulations (0.4 c/deg for CM-tests and 0.2 c/deg for OM-tests). Bandwidths were similar to those of the first stage filters. The outputs of each filter in the quadrature pair were squared and added to produce an 'energy' map for texture modulations at the appropriate frequency. We took the average response across all

image locations and the square root of this value was used as the input to a sigmoidal transfer function given by the equation $r=(bi)^2/(a+(bi)^2)$ where r is the output of the transducer, i its input, a its semi-saturation constant, and b a scaling factor introduced for modeling purposes. This function is square law at low signal levels (thus returning the model output to be an estimate of average energy) but saturates at high signal levels and served to allow our models to simulate within cue 'dipper' functions. We did not include any contrast normalization stage or supra-threshold cross cue masking in our models. We tested our models on continuous OM stimuli with a horizontal dc-orientation.

6.2.2 *Model architectures*. We present results from four models. The architectures (wiring diagrams) of the models were chosen to represent some of the architectures that have been used previously (see Figure 7). We implemented two models to deal with CM as a test cue. The first 'narrow band' model (Fig 7a) had a single first stage channel whose preferred orientation was chosen to produce the biggest differential response between an image with the test cue alone and a noise only image. For dc-orientations of 90° this channel had a preferred orientation at 90°. This model was designed to test Kingdom et al's (2003) off-axis hypothesis with respect to our stimulus configurations, and in the light of our data. It tests the idea that CM detection is mediated by the single first stage channel that is most sensitive to the carrier.

The second 'broad band' model (Fig 7b) had a bank of first stage filters tuned to orientation in the range 0 – 165° degrees at 15° intervals. The outputs from these filters were summed after the non-linearity and before the second stage filters were applied (*cf* Schofield 2000a). This model is broad-band in orientation not spatial frequency. Such a model is more efficient at detecting CM than a narrowly tuned system when the carrier is broad band in orientation as it extracts more of the available texture energy at the first stage. We hypothesized that this facility would be useful in modeling our results when OM pedestals

are large as in this case the carrier energy for the CM-test is distributed across a wide range of orientations.

When OM was the test cue we implemented a narrow band model much as above except that for OM the preferred orientation for the first stage filters was set to 60° (As Fig 7a except for the orientation tuning of the first stage filter). This gives a stronger differential response for weak OM signals than does a channel tuned to the carrier and is similar to the optimal channel orientation found by Kingdom et al (2003). (Note that a channel tuned to 120°would serve equally well). Interestingly the optimal channels for detecting OM are close to 30° from the carrier for OM modulation depths up to about 30° (zero to peak) whereafter the optimal channel follows the peak orientation (Prins and Kingdom, 2002). Thus 30° off-axis filters would be close to optimal for the detection of most of our psychophysical stimuli. Again this model was designed to test the off-axis processing hypothesis in the context of our stimuli and results. For our final model (the opponent model, Fig 7c) we applied the second stage filtering to the signal produced by subtracting the post squaring output of a 60° filter from the post squaring output of a 120° filter (*cf* Landy and Bergen, 1991). This is in principle the same as the narrow band off-axis model except that the former recovers a half-wave rectified version of the modulation signal whereas the latter recovers the orientation modulation more faithfully.

Figure 7 about here

We did not apply the opponent model to CM-test stimuli because the two first stage filters should give equal responses to contrast modulations of the dc-carrier (no OM). Subtracting these signals would thus result in zero output and thus it would be impossible to establish a baseline ‘threshold’ for such a model. Similarly, we did not apply the broad band model to the OM-test case because when OM is the test cue we would expect first stage filters with

opponent orientations to produce equally strong signals of opposing phase which when summed would yield zero response. The channel tuned to the dc-orientation might respond to the OM signal un-opposed but will give only a weak signal.

6.2.3 *Model calibration to psychophysical conditions.* In all of the modelling work we compared the arbitrarily scaled output value from the model for the test cue (with or without pedestal) and compared it to the output for the no test cue case (with or without pedestal). We call this the differential output Δr. We first measured Δr for test cues set to the *average* psychophysical level in the no pedestal case (Δr-baseline). We then measured Δr in the presence of psychophysically matched pedestals (Δr-pedestal) and adjusted the strength of the test cue in order to equate Δr-pedestal with Δr-baseline. The model test cue thresholds thus recorded were directly comparable with human data.

Models were calibrated as follows. Having established Δr-baseline for the test cue we then introduced a same cue pedestal with a range of strengths and adjusted the values of a and b in the final transducer equation to produce a 'classic' within cue dipper function from the modelled thresholds. Model threshold curves had minima when the pedestals were presented at there own, psychophysically determined, detection thresholds (model data not shown).

Having established a set of model parameters that would give clear facilitation for within cue conditions we then fixed these parameters and tested the models in the cross cue conditions. We did not adjust any of the model parameters in order to improve the degree of fit between model and data. We re-measured Δr-baseline and then introduced cross-cue pedestals at psychophysically appropriate levels and in each case adjusted the required strength of the test cue so that Δr-pedestal equalled Δr-baseline. We did this for both phase

relationships. Throughout, our model responses were based on the average response to a number of stimulus examples with different absolute phases and based on different noise samples.

6.2.4 *Model predictions.* Model predictions (that is the relative 'threshold' test cue required to keep the differential response constant with changing pedestal level) are shown on figures 4 and 5. Because we did not optimize the models to individual data sets all curves for a given model and phase relationship are the same. Because model parameters were chosen to best match average tuning curves and calibrated against average threshold values, it is most appropriate to compare the model prediction to the averaged data. Comparisons with individual data are shown for completeness.

We assessed the quality of fit for each model using a Monte Carlo technique similar to that used in section 6.1.2. First we computed best fit quadratic curves for the averaged data in Figs 4 and 5. We then generated simulated data sets based on the averaged data and their standard errors. For each simulated data set we then computed the sum of squared errors for the quadratic fits, null hypothesis (y=1.0) and each of the two applicable models for each condition (10,000 new data sets were generated for each 'model') thus bracketing the models with best and worst case alternatives. We then computed the mean of the sum of squared errors for each 'model' (top half of table 3). We then compared the simulated sum of squared errors for the model predictions with those for the null hypothesis and quadratic fits. The resulting probability estimates for each comparison are shown in the lower half Table 3, these represent the probability with which the null hypothesis or quadratic fit (respectively) can be expected to produce a lower sum of squared error than the models.

Table 3 about here.

The narrow-band model provides a poor match to the data when CM is the test cue, predicting inhibition in the phase relationship B case and (perhaps) facilitation in the phase relationship A case (see fig 4, dotted lines). The sum of squared errors is higher for this model than for the null hypothesis in phase relationship B, and the null hypothesis 'beats' the narrow band model with probability 1.0 (to 4dp) in this phase relationship. Unsurprisingly the quadratic fit gives a better sum of squared error than the model (which was not optimized to the data) but more importantly it also beats the model with probability 1.0 (to 4dp) in the phase relationship B case.

The broad band model (fig 4, dashed lines) provides much better predictions for the CM-test data showing weak facilitation in phase relationship B, and masking in phase relationship A. The level of facilitation although small is consistent with the psychophysical results. The broad band model produces a lower sum of squared errors estimate than the null hypothesis in all conditions. Although the probability estimates for this comparison are quite high the model does tends to beat the null prediction. As should be expected the quadratic fit provides a better overall prediction than the broad band model; the probability estimates shows that the quadratic fit tends to beat the model but this is not significant ($p<0.95$). We can conclude that the broad band model provides a better (and reasonably good) fit to the data than either the narrow band model or the null hypothesis prediction.

Model predictions for OM test cues are shown on Figure 5. Both the narrow band model with off axis processing (dotted line) and the opponent model (dashed line) provide a good qualitative fit to the data showing facilitation in phase relationship B and inhibition / masking in phase relationship A. The opponent model is generally better than the narrow band model (lower sum of squared errors in table 3) and this model has a lower sum of squared error than the null hypothesis. Estimated probabilities (lower part of table 3) show

that the null prediction tends to beat the narrow band model in the phase relationship A case whereas the opponent model tends to do better than the null prediction. The comparison between the models and the null prediction is in favor of the models for the phase relationship B case and this is significant. As should be expected neither model provides a better fit to the data than the quadratic fit but the opponent model compares more favorably to it than the narrow band model. We conclude then that the opponent model provides a better (and reasonably good) fit to the data than the narrow band model. The degree of fit between this model and the data is significantly better than the null hypothesis for the phase relationship B case.

We should not necessarily expect significant results for our models. Given the lack of free parameters, the degree of agreement (quality of fit) between the broad band model in the CM-test case and both narrow band and opponent models in the OM-test case is remarkable.

6.2.5 *Proposal for a generalized texture mechanism*

We envisage our successful models above as being expressions of some generalised texture mechanism. We propose that this mechanism consists of a bank of FRF-like sub-channels. There would need to be many such sub-channels for each modulation spatial frequency and orientation, each sensitive to different carrier properties. These sub-channels might be permanently interconnected or the interconnections could be deployed flexibly. When the carrier signal for the texture modulations is narrow-band separate sub-channels with no real interconnectivity would be adequate to explain our data in the CM-test condition, however, when the background stimulus contains large orientation modulations a broad band model which sums across sub-channels provides a better description of the data. When OM is the test cue, subtractive interconnections between opponent orientation channels (as are

common in the texture segmentation literature) would predict the data more accurately than a single narrow channel.

Above we propose two models for texture detection: a broad band (summing) model for detecting CM and an orientation opponent (subtracting) model for detecting OM. These two models are easily reconciled within a single texture processing system. As discussed in section 6.2.2 neither of these models are very good at detecting the other's preferred cue. Thus instantiations of the two models could co-exist as independent processing streams because they are optimised for different cues. One does not need to posit a mechanism for choosing between the two streams based on the stimulus because the organism might simply respond according to the signals from the most active stream. That the two cue types can sometimes interact within one or other of these streams is evident from the model predictions.

We note that this generalised scheme is very similar to those proposed by Landy and Bergen (1991) and Kingdom et al (2003) except that in our version the output of the broad band CM mechanism must be made available to future processing stages. Our data do not imply a role for the broad-band CM channel as part of a contrast normalization process when CM is presented at twice the frequency of OM but they do not exclude the use of the CM mechanism to provide a normalisation signal for the OM mechanism under other conditions.

7 General discussion

The facilitatory trends in our data are relatively weak, however they do have statistical support and, perhaps more importantly, they can be modeled reasonably well with FRF sub-channels that have no truly free parameters. The facilitation of OM at 0.4c/deg by CM at 0.2c/deg in phase relationship A is statistically significant. Facilitation in the opposite

direction (CM-test) is weaker but still significant and reproduced well by our preferred model. Facilitation depends on the relative phase of the cues and starts to act when the pedestal cue is at its own threshold level. Given the phase dependence and the fact that we applied positional jitter within our display intervals we are confident that the facilitation observed is not due to positional cues. We find threshold minima when the pedestal is between 2 and 4 times its own threshold. This is a little late, however weak facilitation might imply only weak cross-talk between cues and this might shift the threshold minima as the pedestal cue would have to be greater than its own threshold to exert one 'threshold's worth' of influence on the other cue.

We found facilitation only in phase relationship B; that is when the peaks of contrast align with peaks and troughs of orientation. When OM is the test cue the introduction of a CM pedestal will boost the carrier signal at the orientation extremes (peaks and troughs in orientation) and in the limit the signal will begin to look like strips of high contrast Gabor noise with alternating orientation separated by regions of low contrast. This effect can be seen in see figure 3a and the icons below this image indicate which orientations are preserved. Such stimuli might provide a good cue to OM sensitive mechanisms. In phase relationship A where the high contrast regions align with the orientation zero-crossings, the OM signal is attenuated at the orientation extremes (peaks and troughs) while regions of dc-orientation are boosted and in the limit signals will begin to look like strips of Gabor noise all with the same orientation (see figure 3d and associated icons). This would be a relatively poor stimulus for OM detectors.

It is harder to see why facilitation should take place in phase relationship B when CM is the test cue. Indeed our narrow band model suggests that inhibition should occur in this condition as the OM will induce small fluctuations in a channel tuned to the dc-orientation that are in phase with CM only when the two cues are in phase relationship A. However as

the OM pedestal strength increases the carrier signal for the CM-test becomes more broad band in orientation. In such a circumstance a narrow band channel will waste some of the available carrier energy by filtering it out at the first stage. A broadly tuned first stage channel is more efficient in this respect. We propose that if the carrier is orienationally broad-band CM is detected within a broad-band mechanism constructed from narrow band sub-channels. Some of these sub-channels will be tuned to the extremes of orientation in the OM pedestal and their output will vary in response to both OM and CM. However, the two influences will now be in phase when the cues are in phase relationship B not A.

The suggestion that CM is processed in a broad-band channel is at odds with the results of Dakin and Mareschal (2000) who found that noise masks CM detection in a way that is consistent with narrowly tuned first-stage filters. However in their study the carrier signals were narrow band, and in this case a narrowly tuned first stage would help to exclude the influence of additional un-modulated noise in the stimulus. Our strong OM pedestals represent a broad-band carrier and in this case broad-band processing may be more efficient. Flexible wiring of CM sub-channels could reconcile our results and those of Dakin and Mareschal (2000) allowing the CM mechanisms to optimize itself to the properties of its carrier.

Our result are also somewhat at odds with those of Kingdom et al (2003). We found weak facilitatory interactions whereas they found no interactions at all. We note however that we only find facilitation in conditions that Kingdom et al (2003) did not test. They only tested OM and frequency-doubled CM in phase relationship A. Like them we found little interaction in this case. Further although they tested non frequency-doubled combinations in several phase relationships they did so only for pedestals at their own detection threshold. Our facilitation is very weak at such pedestal levels. Further, Kingdom et al (2003) only tested with high pedestal strengths in the non frequency-doubled case for one

phase relationship. It is not clear which phase relationship they used but if it were the equivalent of our phase A then our results would be consistent with theirs. Even if they used the equivalent of our phase relationship B high contrast regions in their stimuli would have aligned with alternate orientation extremes (peaks *only* or troughs *only*) and these points in the image would have the same orientation (not opposite orientations as in our case). Thus, in the limit, strong CM would tend to render their stimuli as stripes of Gabor noise with the same orientation and it is not clear that this would support improved OM detection. In our frequency-doubled phase B case, strong CM renders the image as alternating strips of oppositely oriented Gabor noise which should be a good cue for OM detection. Finally, we have previously found an absence of facilitation for non frequency-doubled interactions in both phase relationships (Schofield 2000b). Thus we conclude that our results are broadly consistent with those of Kingdom et al (2003) where our experimental conditions overlap but we have tested conditions that are perhaps of more diagnostic utility.

Our results and modeling work provide positive support for the argument put forward by Kingdom et al (2003) that there is a single generalized texture mechanism with many sub-channels and that the processing of different cues about the same carrier takes place in separate sub-channels due to off-axis processing. Crucially, even in this off-axis model some interaction is predicted unless a contrast normalization stage is added and we now present evidence for that interaction. The models proposed by Kingdom et al (2003) and Landy and Bergen (1991) would not show facilitation of OM by CM because their contrast normalization stages (proposed but not implemented in the case of Kingdom et al, 2003) would effectively remove the CM component. Our results suggest that contrast normalization of OM signals if it exists at all does not take place when CM pedestals are at twice the frequency of the OM tests.

We found few differences based on the percept formed by the OM signal. We did, however, find some evidence that late masking of CM-tests by OM may depend on the perceptual properties of the OM cue (a result we have noted before, Schofield 2000b). Masking is strongest when OM forms a segmented percept. This may be due to the introduction of texture boundaries in this cue that are not present in the continuous case. This phenomenon warrants further study.

Acknowledgements

This work was funded by an EPSRC project grant (GR/R17249/01) to AJS.

References

Cavanagh P, Mather G, 1989 "Motion: The long and short of it" *Spatial Vision* **4** 103-129

Chubb C, Sperling G, 1988 "Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception" *Journal of the Optical Society of America A* **5** 1986-2007

Dakin S, Mareschal I, 2000 "Sensitivity to contrast modulation depends on carrier spatial frequency and orientation" *Vision Research* **40** 311-329

Derrington A. M, Badcock D. R, Henning G. B, 1993 "Discriminating the direction of second-order motion at short stimulus durations" *Vision Reseach* **33** 1785-1794

DeValois R. L, Albrecht D. G, Thorell L. G, 1982 "Spatial-frequency selectivity of cells in macaque visual cortext" *Vision Research* **22** 545-559

Foster D. H, Bischof W. F, 1991 "Thresholds from psychometric functions: superiority of bootstrap to incremental and probit variance estimators" *Psychological Bulletin* **109** 152-159

Georgeson M. A, Schofield A. J, 2002 “Shading and texture: separate information channels with a common adaptation mechanism?” *Spatial Vision* **16** 59-76

Gray R, Regan D, 1998 “Spatial frequency discrimination and detection characteristics for gratings defined by orientation texture” *Vision Research* **38** 2601-2617

Howell D.C, 1997 *Statistical methods for psychology 4th Ed*. Duxbury Press, Belmont CA.

Keeble D. R. T, Kingdom F. A. A, Morgan M. J, 1997 “The orientation resolution of Human Texture Perception” *Vision Research* **37** 2993-3007

Keeble D. R. T, Kingdom F. A. A, Moulden B., Morgan M. J, 1995. “Detection of orientationally multimodal textures” *Vision Research* **35** 1991-2005

Kingdom F. A. A, Keeble D, Moulden B, 1995 “Sensitivity to orientation modulation in micropattern-based textures” *Vision Research* **35** 79-91

Kingdom F. A. A, Prins N, Hayes A, 2003 “Mechanism independence for texture-modulation detection is consistent with a filter-rectify-filter mechanism” *Visual Neuroscience* **2** 65-76

Klien S.A, Hu Q.J, Carney T, 1996 “The adjacent pixel nonlinearity: problems and solutions” *Vision research* **36** 3167-3181

Landy M. S, Bergen J. R, 1991 “Texture segregation and orientation gradient” *Vision Research* **31** 679-691

Ledgeway T, 1994 “Adaptation to 2nd-order motion results in a motion aftereffect for directionally-ambiguous test stimuli” *Vision Research* **34** 2879-2889

Ledgeway T, Smith A. T, 1994 “Evidence for separate motion-detecting mechanisms for first-order and 2nd-order motion in human vision” *Vision Research* **34** 2727-2740

Ledgeway T, Smith, A. T, 1995 “The perceived speed of 2nd-order motion and its dependence on stimulus contrast” *Vision Research* **35** 1421-1434

Legge G. E, Foley J.M. 1980 “Contrast masking in human vision” *Journal of the Optical Society of America* **70** 1458-1471

Lu Z. L, Sperling G, 1996 “Contrast gain control in first- and second-order motion perception” *Journal of the Optical Society of America A* **13** 2305-2318

Luschow A, Nothdurft H. C, 1993 “Pop-out of orientation but no pop-out of motion at isoluminance” *Vision Research* **33** 91-104

Malik J, Perona P, 1990 “Preattentive texture processing with early vision mechanisms” *Journal of the Optical Society of America A* **7** 923-932

Motoyoshi I, Kingdom F. A. A, 2003 “Orientation opponency in human vision revealed by energy-frequency analysis” *Vision Research* **43** 2197-2205

Mulligan J.N, Stone L.S, 1989 “Halftoning method for the generation of motion stimuli” *Journal of the Optical Society of America A* **6** 1217-1227

Nothdurft H. C, 1985 “Orientation sensitivity and texture segmentation in patterns with different line orientation” *Vision Research* **25** 551-560

Nothdurft H. C, 1991 “Texture segmentation and pop-out from orientation contrast” *Vision Research* **31** 1073-1078

Nothdurft H. C, 1993 “The role of features in preattentive vision - comparison of orientation, motion and color cues” *Vision Research* **33** 1937-1958

Prins N, Kingdom F. A, 2002 “Orientation- and frequency-modulated textures at low depths of modulation are processed by off-orientation and off-frequency texture mechanisms” *Vision Research* **42** 705-713

Schofield A. J, 2000a “What does second-order vision see in an image?” *Perception.* **29** 1071-1086

Schofield A.J, 2000b “Interactions between orientation and contrast in the processing of texture cues” *Perception **29** Supplement*, p50

Schofield A. J, Foster D. H, 1995 “Artificial neural networks simulating visual texture segmentation and target detection in line-element images” *Philosophical Transactions of the Royal Society (London) B* **350** 401-412

Schofield A. J, Georgeson M. A 1999 “Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behaviour” *Vision Research* **39** 2697-2716

Schofield A. J, Georgeson M. A, 2000 “The temporal properties of first- and second-order vision” *Vision Research* **40** 2475-2487

Schofield A. J, Georgeson M. A 2003 “Sensitivity to contrast modulations: the spatial frequency dependence of second-order vision” *Vision Research* **43** 243-259

Scott-Samuel N. E, Georgeson M. A 1999 "Does early nonlinearity account for second-order motion?" *Vision Research* **39** 2853-2865

Smith A. T, Hess R. F, C.L.Baker Jr, 1994 "Direction-identification thresholds for second-order motion stimuli" *Journal of the Optical Society of America A* **11** 506-514

Smith A. T, Ledgeway T, 1997 "Separate detection of moving luminance and contrast modulations: fact or artefact?" *Vision Research* **37** 45-62

Snowden R.J, 1992 "Orientation bandwidth: the effect of spatial and temporal frequency" *Vision Research* **32** 1965-1974

Solomon J. A, Sperling G, 1994 "Full-wave and half-wave rectification in second-order motion perception" *Vision Research* **34** 2239-2257

Solomon J. A, Sperling G, 1995 "1st-order and 2nd-order motion and texture resolution in central-vision and peripheral-vision" *Vision Research* **35** 59-64

Sutter A, Sperling G, Chubb C, 1995 "Measuring the spatial-frequency selectivity of 2nd-order texture mechanisms" *Vision Research* **35** 915-924

Wichmann F. A, Hill N. J, 2001a "The psychometric function: I. Fitting, sampling and goodness-of-fit" *Perception and Psychophysics* **63** 1293-1313

Wichmann F. A, Hill N. J, 2001b "The psychometric function: II. Bootstrap-based confidence intervals and sampling" *Perception and Psychophysics* **63** 1314-1329

Wilson H. R, Ferrera V. P, Yo C, 1992 “A psychophysically motivated model for two-dimensional motion perception” *Visual Neuroscience* **9** 79-97

Figure legends

Figure 1. Orientation-modulated anisotropic noise. (a) Continuous appearance: baseline or dc-orientation is horizontal, modulation orientation is vertical. (b) Cross section through the modulation waveform which defines the orientation at any point in (a). (c) Segmented appearance: dc-orientation is vertical, modulation orientation also vertical. (d) Cross section of the modulator for (c). Here and elsewhere images are extracts (central ¼ area) from experimental stimuli rendered for publication. Vertical lines link peaks, troughs and zero-crossings in the modulation waveform to corresponding points in the image. Line icons below the image are an approximate representation of the dominant orientation at the key points.

Figure 2. Contrast-modulated uniformly oriented anisotropic noise. (a) Extract from an experimental stimulus with 100% contrast modulation depth. (b) Cross section through the modulation waveform for (a). Vertical lines link contrast peaks and troughs in the modulator to corresponding points in the stimulus. ‘H’ and ‘L’ labels below the image indicate high and low contrast respectively. The dc-orientation (horizontal) is most visible at the contrast peaks which occur at the same frequency as the zero crossings in Figure 1b.

Figure 3. Combinations of OM and CM. (a) Continuous type OM with CM in phase relationship B (peaks of contrast align with peaks and troughs of orientation – that is, orientation extremes). (b) Cross section through the OM modulation waveform for (a). (c) cross section through the CM modulation waveform for (a). (d) Segmented type OM with CM in phase relationship A (peaks of contrast align with zero crossings). (e) Cross section through the OM modulation waveform for (d). (f) cross section through the CM modulation waveform for (d). Vertical lines in (b) and (e) mark the peaks, troughs and zero-crossings of orientation, similar lines in (c) and (f) mark the peaks and troughs of contrast. ‘H’ and ‘L’

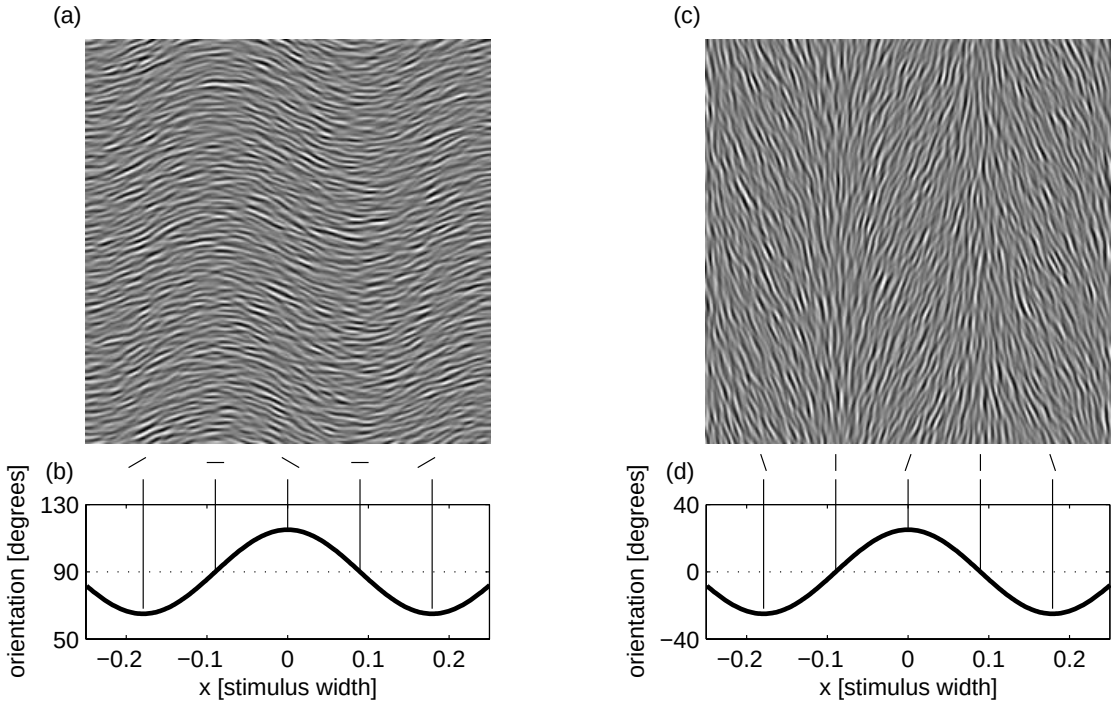
labels indicate points of high and low contrast respectively. Line icons below each image indicate orientations at key points in the image that survive the contrast modulation process.

Figure 4. Results of experiment 1: CM on OM. Thresholds for detecting a CM test cue (scaled relative to the no pedestal CM baseline condition) are plotted as a function of the strength of the OM pedestal scaled relative to its own detection threshold. (a-d) Data from all cases where the OM and CM were in phase relationship A. (e-h) Data for phase relationship B. Open symbols with short ticks on error bars represent data for the segmented conditions, filled symbols with long ticks on error bars data for continuous percepts. Panels a & e show data for observer AJS, b & f for AGC, and c & g for TAY). Panels d & h show data averaged across all observers and both percepts (solid lines and squares). Dashed lines are predictions from a broad-band model, dotted lines are predictions from a narrow band model (see Section 6.2). Models had no free parameters and were not optimized to the data; hence model curves are the same in all panels. Error bars show 95% confidence intervals but see text.

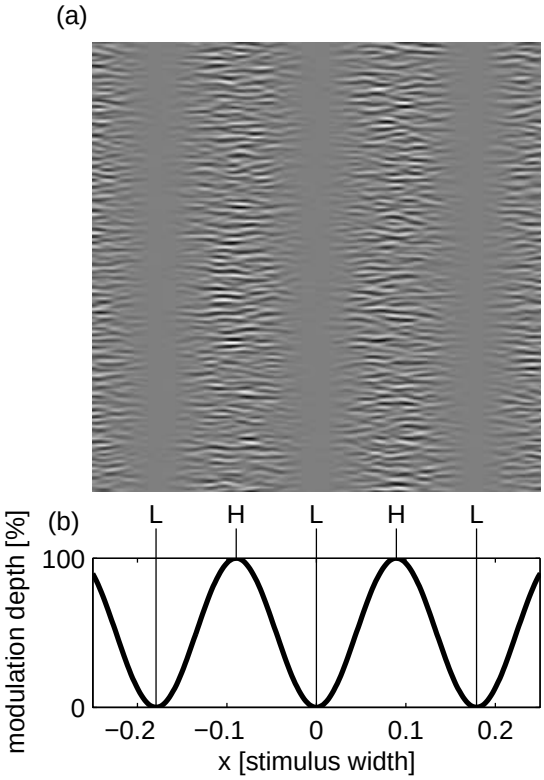
Figure 5. Results of experiment 2: OM on CM. Thresholds for detecting an OM test cue (scaled relative to the no pedestal OM baseline condition) are plotted as a function of the strength of the CM pedestal scaled relative to its own detection threshold. (a-c) Data from all cases where the OM and CM were in phase relationship A. (d-f) Data for phase relationship B. Symbols are as for Fig 4. Panels a & d show data for observer AJS, and b & e for TAY). Panels c & f show data averaged across all observers and both percepts (solid lines and squares). Dashed lines are predictions from the opponent orientation model, dotted lines are predictions from the narrow band model (see Section 6.2). Models had no free parameters and were not optimized to the data; hence model curves are the same in all panels. Error bars show 95% confidence intervals but see text.

Figure 6. OM on OM control. Relative thresholds for detecting an OM test cue as a function of the relative detection threshold for the OM pedestal for observer TAY (note log-log axes). Error bars are 95% confidence intervals estimated by bootstrapping. (a) Segmented appearance, (b) continuous appearance. Test and pedestal were always in phase.

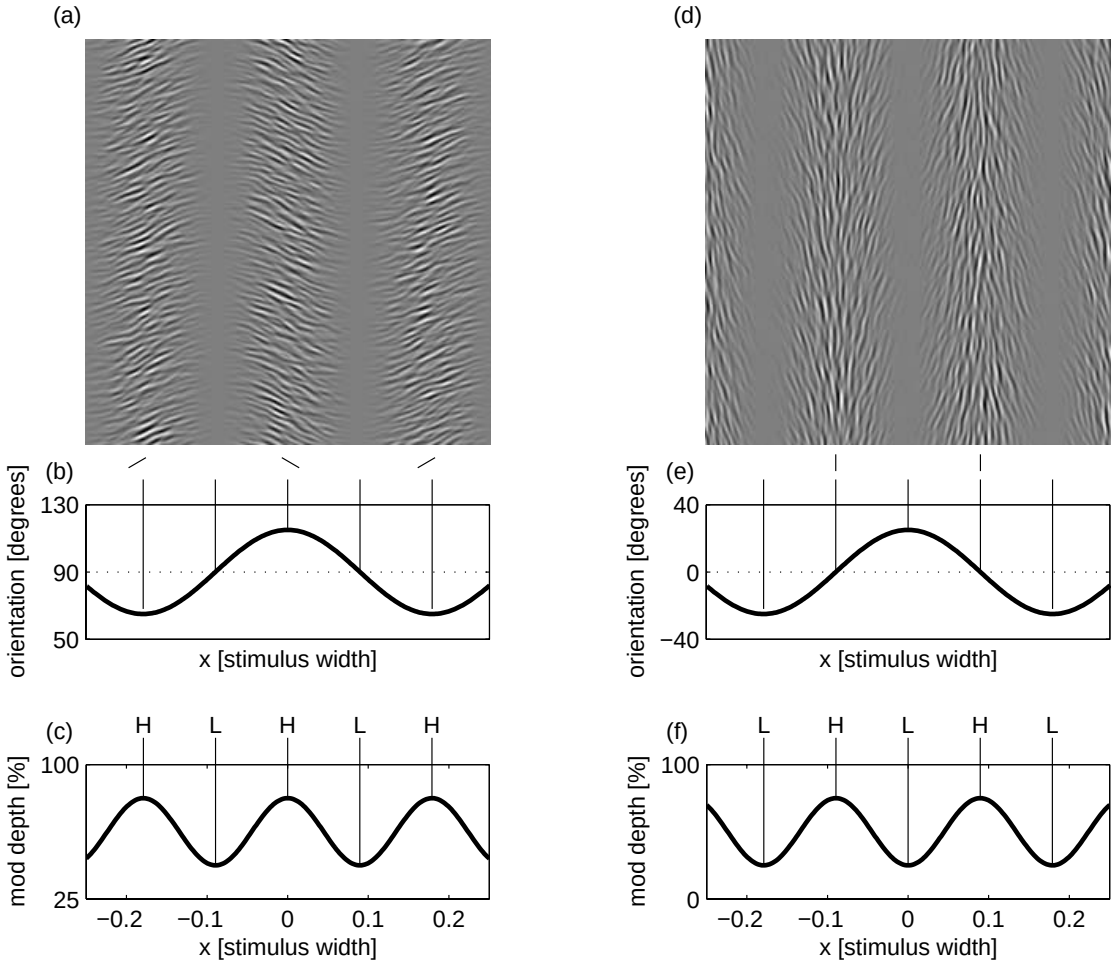
Figure 7. Schematic representations of the model sub-channels used to generate the predictions of figures 4 and 5. a) Narrow band model as used when CM was the test cue. b) broad-band model used when CM was the test. c) Opponent model used when OM was the test. The narrow band model used when OM was the test was as panel (a) except for the orientation of the first stage filters (see text).



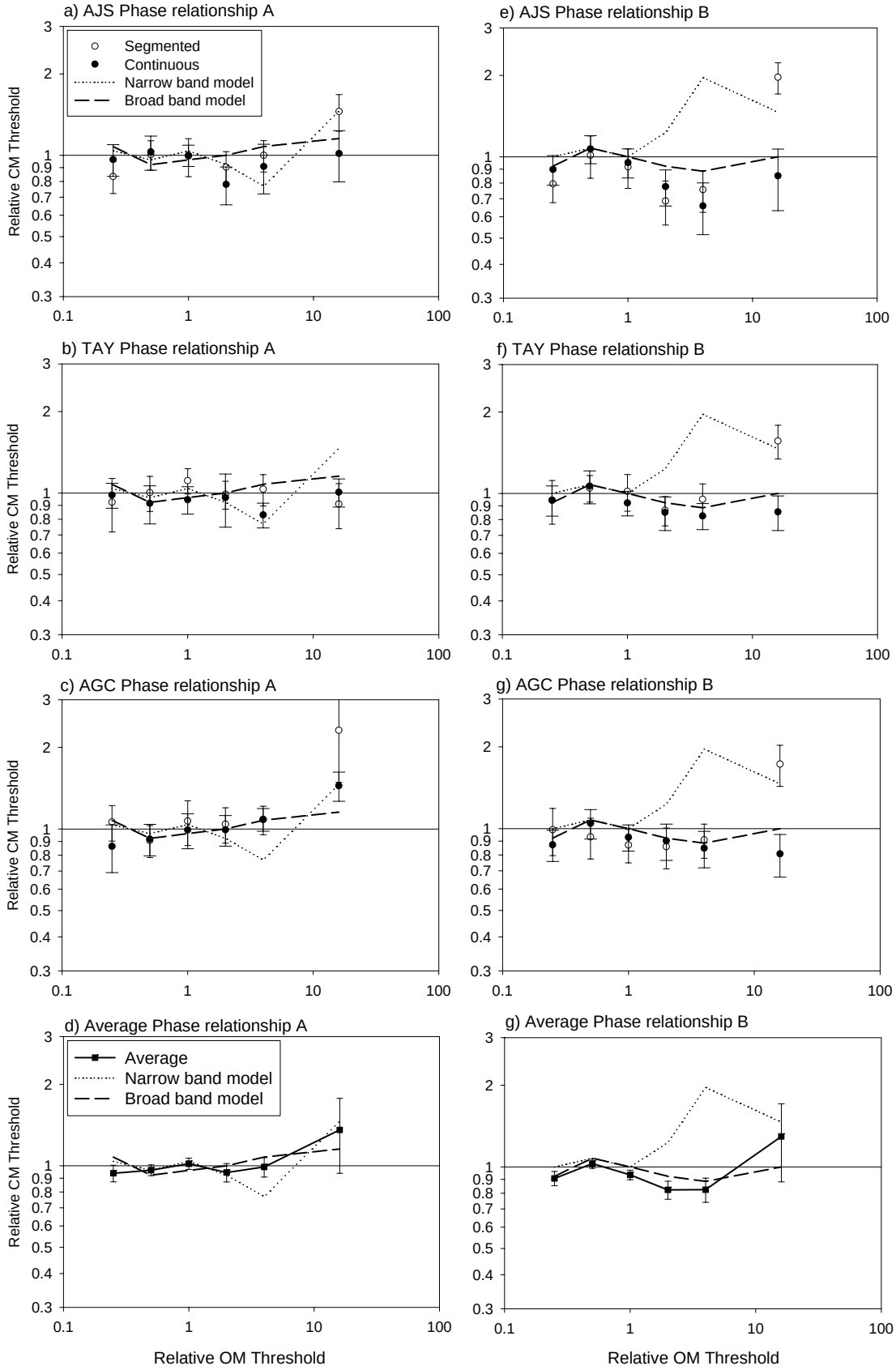
Schofield and Yates Figure 1



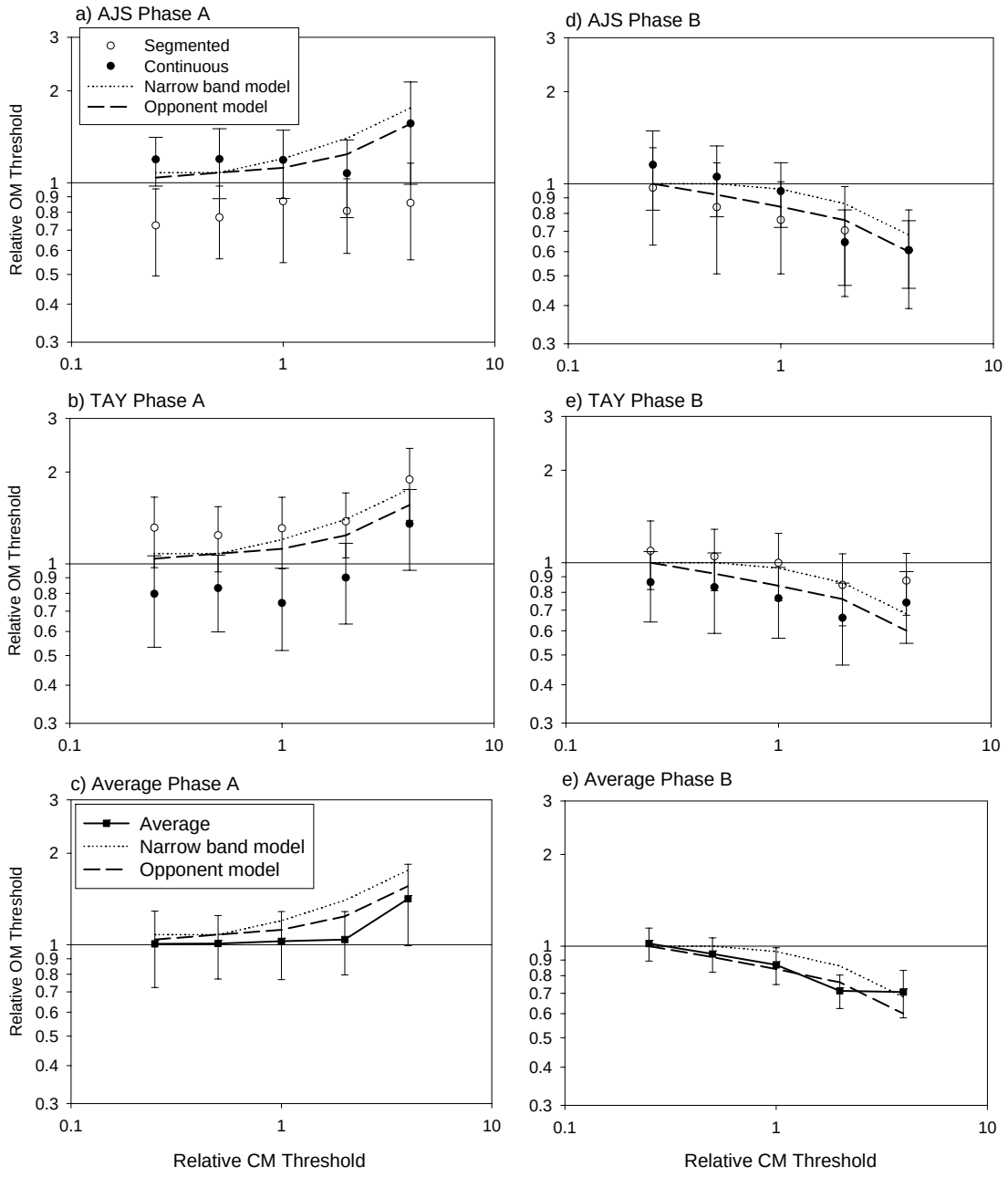
Schofield and Yates, Figure 2



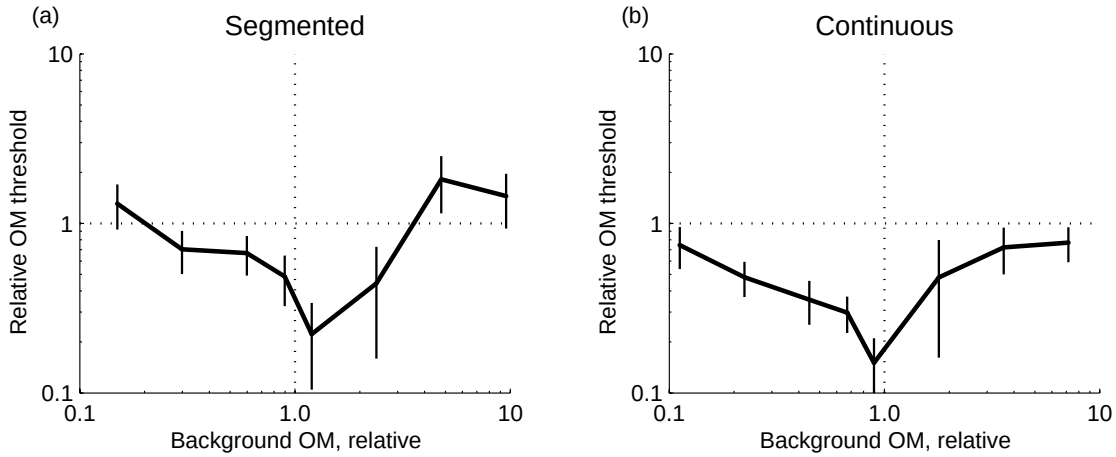
Schofield and Yates, Figure 3



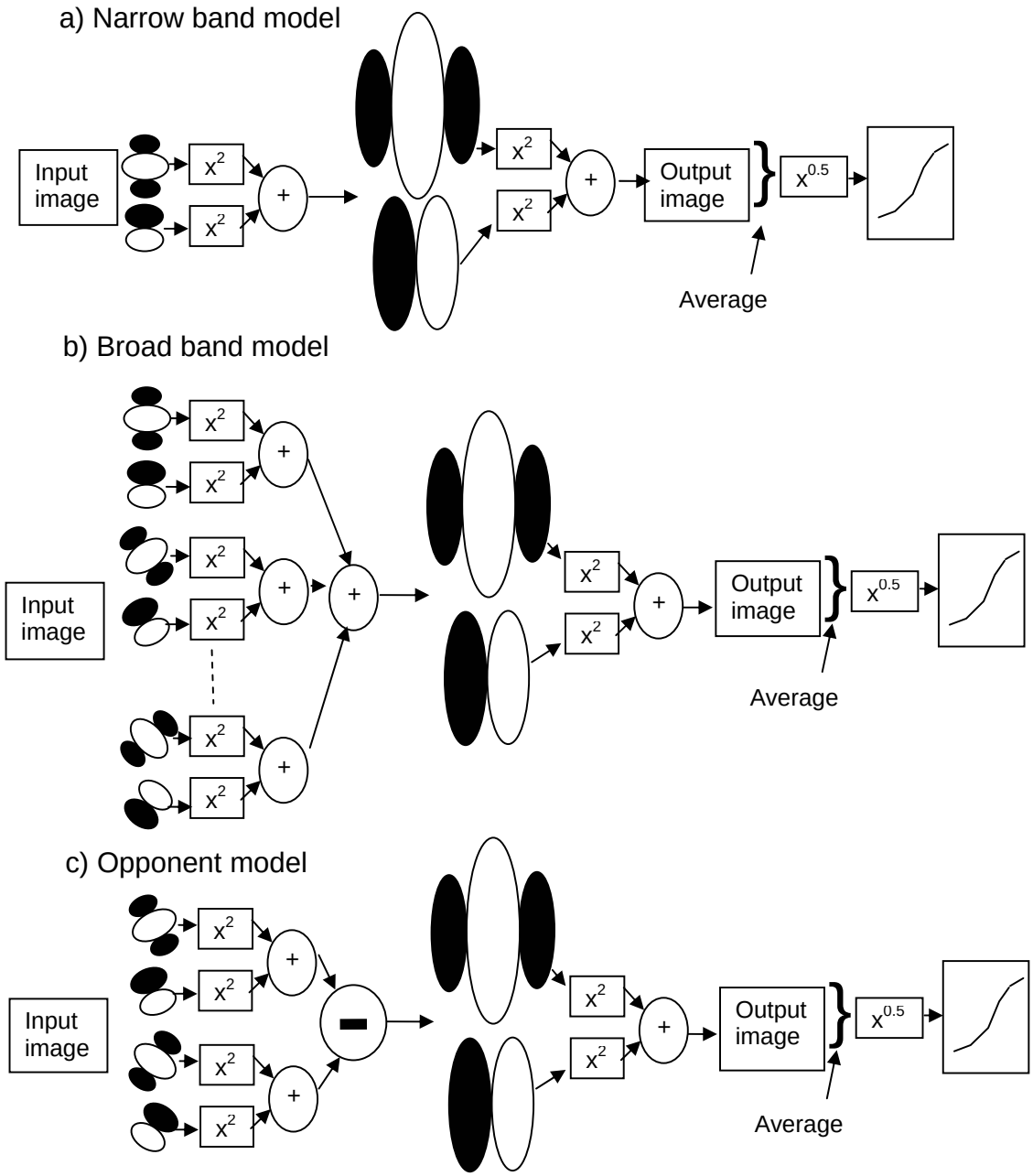
Schofield and Yates Figure 4



Schofield and Yates, Figure 5



Schofield and Yates, Figure 6



Schofield and Yates, Figure 7

Table 1. Baseline detection thresholds for OM and CM in the absence of a pedestal cue used for the normalization of the x-axes in figures 4 and 5.

	OM thresholds [Δ° zero to peak]				CM thresholds [modulation depth]			
	Vertical carrier		Horizontal carrier		Vertical carrier		Horizontal carrier	
Observer	Vertical mod	Horizontal mod	Vertical mod	Horizontal mod	Vertical mod	Horizontal mod	Vertical mod	Horizontal mod
TAY	1.75	1.85	1.34	1.52	0.11	0.14	0.12	0.09
AJS	4.22	3.21	3.31	4.17	0.15	0.19	0.16	0.13
AGC	1.97	1.75	1.82	2.61	0.10	0.12	0.11	0.11

Table 2. Estimates of the probability that a model representing the null hypothesis will produce a lower sum of squared error than the best fitting quadratic curve for individual data as estimated by Monte Carlo simulation.

	CM test on OM pedestal				OM test on CM pedestal			
	Phase relationship A		Phase relationship B		Phase relationship A		Phase relationship B	
Observer	Seg	Cont	Seg	Cont	Seg	Cont	Seg	Cont
TAY	0.4120	0.1556	0.0009	0.0295	0.0025	0.0641	0.2713	0.0038
AJS	0.0075	0.2694	0.0000	0.0036	0.0642	0.0986	0.026	0.0075
AGC	0.0276	0.0008	0.0005	0.0488				

Table 3. Sum of squared error comparisons for the computational models. Top half of table: The mean of the sum of squared errors for each model, the best quadratic fit, and the null hypothesis. Bottom half of table: Columns headed ‘vs null’ show the estimated probability that the null hypothesis produces a lower sum of squared errors than would predictions for a given model. Columns headed ‘vs quad’ show the estimated probability that the best fitting quadratic curve to the averaged data would produce a lower sum of squared errors than the predictions for a given model. All values are estimates based on Monte Carlo simulations for the averaged data for Figure 4 and 5.

a) CM test on OM pedestal

	Sum of squared errors			
Model	Phase rel’ A		Phase rel’ B	
Null	0.1864		0.2121	
Quad	0.0546		0.0614	
Narrow	0.1201		1.5511	
Broad	0.1262		0.1584	
	Probability estimates			
	Phase rel’ A		Phase rel’ B	
Model	vs null	vs quad	vs null	vs quad
Narrow	0.4184	0.8259	1.0000	1.0000
Broad	0.4089	0.7894	0.3437	0.7638

b) OM test on CM Pedestal

	Sum of squared errors			
Model	Phase rel’ A		Phase rel’ B	
Null	0.2947		0.2078	
Quad	0.1161		0.0182	
Narrow	0.4032		0.0522	
Opponent	0.1912		0.0329	
	Probability estimates			
	Phase rel’ A		Phase rel’ B	
Model	vs null	vs quad	vs null	vs quad
Narrow	0.6757	0.9396	0.0011	0.9226
Opponent	0.3453	0.7103	0.0007	0.7496